The PHQ-9: A New Depression Diagnostic and Severity Measure

Kurt Kroenke, MD; and Robert L. Spitzer, MD

Depression is one of the most prevalent and treatable mental disorders presenting in general medical as well as specialty settings. There are a number of case-finding instruments for detecting depression in primary care, ranging from 2 to 28 items.[1,2] Typically these can be scored as continuous measures of depression severity and also have established cutpoints above which the probability of major depression is substantially increased. Scores on these various measures tend to be highly correlated[3], with little evidence that one measure is superior to any other.[1,2,4]

PHQ AND PHQ-9

The primary care evaluation of mental disorders (PRIME-MD®) is a novel instrument developed a decade ago to assist primary care clinicians in making criteria-based diagnoses of five types of *DSM-IV* disorders commonly encountered in medical patients: mood, anxiety, somatoform, alcohol, and eating.[5,6] The patient health questionnaire (PHQ) is a 3-page self-administered version of the PRIME-MD® that has been well validated in two large studies involving 3,000 patients in 8 primary care clinics and 3,000 patients in 7 obstetrics–gynecology clinics.[7,8] Because it is entirely self-administered and has diagnostic validity comparable to the clinician-administered PRIME-MD®, the PHQ is now the most commonly used version in both clinical and research settings.

At 9 items, the PHQ depression scale (which we call the PHQ-9) is half the length of many other depression measures, has comparable sensitivity and specificity, and consists of the actual nine criteria on which the diagnosis of *DSM-IV* depressive disorders is based.[9] The latter feature distinguishes the PHQ-9 from other two-step depression measures for which, when scores are high, additional questions must be asked to establish *DSM-IV* depressive diagnoses. The PHQ-9 is thus a dual-purpose instrument that, with the same nine items, can establish provisional depressive disorder diagnoses as well as grade depressive symptom severity.

An item was also added to the end of the diagnostic portion of the PHQ-9 asking patients who checked off any problems on the questionnaire: "How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?" This single item is an excellent global rating of functional impairment and has been shown to correlate strongly with a number of quality of life, functional sta-

Dr. Kroenke is from the Department of Medicine, Regenstrief Institute for Health Care and Indiana University School of Medicine, Indianapolis, Indiana. Dr. Spitzer is from the New York State Psychiatric Institute and Department of Psychiatry, Columbia University, New York, New York. Address reprint requests to Kurt Kroenke, MD, Regenstrief Institute for Health Care, RG-6, 1050 Wishard Blvd, Indianapolis, IN 46202.

Dr. Kroenke's research is supported by Pfizer Inc. and Eli Lilly. He is also a member of Eli Lilly's advisory board. Dr. Spitzer has no industry relationships to disclose.

TABLE 1

PHQ-9 Scores and Proposed Treatment Actions

PHQ-9 Score	Depression Severity	Proposed Treatment Actions
1 to 4	None	None
5 to 9	Mild	Watchful waiting; repeat PHQ-9 at follow-up
10 to 14	Moderate	Treatment plan, considering counseling, follow-up and/or pharmacotherapy
15 to 19	Moderately Severe	Immediate initiation of pharmacotherapy and/or psychotherapy
20 to 27	Severe	Immediate initiation of pharmacotherapy and, if severe impairment or poor response to therapy, expedited referral to a mental health specialist for psychotherapy and/or collaborative management

tus, and health care usage measures.

PHQ-9 as a Diagnostic Measure

The PHQ-9 is the 9-item depression module from the full PHQ (Sidebar, page xx). Major depression is diagnosed if five or more of the nine depressive symptom criteria have been present at least "more than half the days" in the past 2 weeks, and one of the symptoms is depressed mood or anhedonia. One of the nine symptom criteria ("thoughts that you would be better off dead or of hurting yourself in some way") counts if present at all, regardless of duration. Because the PHQ response set was expanded from the simple "yes/no" in the original PRIME-MD® to four frequency levels, we found that lowering the PHQ threshold from "nearly every day" to "more than half the days" provided the optimal sensitivity and specificity. Other depression is diagnosed if two, three, or four depressive symptoms have been present at least "more than half the days" in the past 2 weeks, and one of the symptoms is depressed mood or anhedonia. As with the original PRIME-MD®, before making a clinical diagnosis of a depressive disorder, the clinician is expected to rule out physical causes of depression, normal bereavement and history of a manic episode.

PHQ-9 as a Measure of Depression Severity

As a severity measure, the PHQ-9 score ranges from 0 to 27, because each of the 9 items can be scored from 0 ("not at all") to 3 ("nearly every day"). Easy-to-remember cutpoints of 5, 10, 15, and 20 represent the thresholds for mild, moderate, moderately severe, and severe depression, respectively.[9] If a single screening cutpoint were to be chosen, we currently recommend a PHQ-9 score of 10 or greater, which has a sensitivity for major depression of 88%, a specificity of 88%, and a positive likelihood ratio of 7.1. The latter means that primary care patients with major depression are seven times more likely to have a PHQ-9 score of 10 or greater than patients without major depression. Suggested treatment actions in response to these various levels of PHQ-9 depression severity are shown in Table 1.

Scores less than 10 seldom occur in individuals with major depression whereas scores of 15 or greater usually signify the presence of major depression.[9] In the gray zone of 10 to 14, increasing PHQ-9 scores are associated, as expected, with increasing specificity and declining sensitivity. The operating characteristics of the PHQ-9 compare favorably to nine other case-finding instruments for depression in primary care, which have an overall sensitivity of 84%, a specificity of 72%, and a positive likelihood ratio of 2.9.[1]

Berwick et al.[10] used receiver operating characteristic (ROC) analysis to determine how several brief mental health measures discriminated between patients with and without major depression. In their study, the area under the curve was 0.89 for the 5-item RAND Mental Health Inventory (MHI), 0.90 for the 18-item MHI, 0.89 for the 30-item general health questionnaire, and 0.80 for the 28-item somatic symptom inventory. In the PHQ study, the area under the curve for major depression was 0.95 for the PHQ-9 and 0.93 for the 5-item MHI.[9] It is unlikely that other

depression-specific measures would be significantly better than the PHQ-9 since an area under the curve of 1.0 represents a perfect test.

OUTCOME MEASURES USED TO EVALUATE DEPRESSION TREATMENT RESPONSE

A particularly important characteristic of a severity measure is its sensitivity to change throughout time. In other words, how precisely do declining or rising scores on the measure reflect improving or worsening depression in response to effective therapy or natural history? Although an exhaustive review of depression measures is beyond the scope of this article (but can be found elsewhere[4,11])a brief discussion of selected measures is warranted. The Hamilton Rating Scale for Depression (HAM-D) has been the criterion standard outcome measure in clinical trials, but it can require 15 to 30 minutes of clinician time to administer and is therefore not feasible in many practice settings. The HAM-D also is also rather complicated to score and requires substantial training to get reasonable interrater agreement. The Montgomery-Asberg Depression Rating Scale is about half as long as the HAM-D and probably just as sensitive to change.[12,13] Like the HAM-D, however, the Montgomery-Asberg scale must be administered by a clinician with special training and is moderately time-intensive. Several self-administered scales—the 21-item Beck Depression Inventory and the 20-item Zung Self-Rating Depression Scale—also have been used as outcome measures but may be somewhat less sensitive to change than the HAM-D.[14] The Symptom Checklist-20 (SCL-20) has been used as an outcome measure in primary care clinical trials,[15-17] although published evidence on its sensitivity to change as well as other psychometric characteristics is limited. Epidemiological and clinical studies have established the 20-item Center for Epidemiologic Studies Depression Scale (CES-D) as a valid measure for identifying depression, but there is less information regarding its sensitivity to change.

PHQ-9 As An Outcome Measure: Preliminary Data On Its Sensitivity To Change

The cross-sectional data from 6,000 patients in the two PHQ validation studies have been instrumental in establishing the criterion and construct validity of the PHQ-9 as both a diagnostic and severity measure. However, data from treatment trials are necessary to determine the utility of the PHQ-9 in monitoring a patient's response to depression treatment. Although definitive evidence is awaiting the completion of several large clinical trials, there is preliminary data from the IMPACT study on the PHQ-9's sensitivity to change when used as an outcome measure. The IMPACT study is a multi-center randomized clinical trial in which more than 1,800 older adults with major depression or dysthymia were randomized to depression case management versus usual care.[18] In the IMPACT study, the primary outcome measure was the SCL-20, a depression severity scale that has been extensively used in depression treatment trials in primary care.[15,16,19]

We examined approximately 150 intervention patients in the IMPACT trial, who at baseline had concurrent (ie, within 7 days of each another) PHQ-9 and SCL-20 scores, and a similar sample of patients who had concurrent scores after 3 months of treatment. The correlation of the two measures at baseline was 0.46 and after 3 months of treatment, 0.63. There were approximately 100 patients who had both concurrent baseline and 3-month scores. The mean decline in the SCL-20 was 0.48, which is an effect size (ie, expressed as number of standard deviations) of 0.71. The mean decline in the PHQ-9 was 6.9, which is an effect size of 0.91. The correlation between SCL-20 and PHQ-9 change scores was 0.50. Effect sizes of 0.5 and 0.8 are typically considered to represent moderate and large changes, respectively.[20] Thus, the change in PHQ-9 scores with depression treatment is similar or greater than the change in SCL-20 scores.

Limitations of this preliminary data should be noted. In addition to the small sample size, there were differences in mode of administration. The SCL-20 was administered using a structured computer interview by a research assistant blinded to treatment group. In contrast, the PHQ-9 was administered by the nurses who also were treating the patients, which could introduce a bias toward greater improvement. When the IMPACT trial is complete, analysis of the full 1,800 patients at multiple time points and including depression

TABLE 2

Comparison of the Likelihood Ratios for Different Levels of PHQ-8 and PHQ-9 Severity Scores in Diagnosing Any Depressive Disorder

	Percent With Any Depression		Likelihood Ratio*	
PHQ Severity Score	**PHQ-8**	**PHQ-9**	**PHQ-8**	**PHQ-9**
1 to 4	0.2%	0.1%	0.010	0.006
5 to 9	12.9%	12.6%	0.79	0.76
10 to 14	58.0%	54.9%	7.3	6.5
15 to 19	91.5%	90.6%	57.0	51.3
20 to 27	98.9%	97.5%	475.5	203.8

**The likelihood ratio expresses how much more likely it is for a patient with a certain disease to have a given test result than it is for a person without the disease. Thus, a patient with any depressive disorder is 7.3 times more likely to have a PHQ-8 score in the10 to 14 range than a person with no depressive disorder.*

diagnostic status as assessed by the SCID will be necessary to confirm the sensitivity of the PHQ-9 to change.

Currently, we consider that a decline in the PHQ-9 score of at least 5 points is necessary to qualify as a clinically significant response to depression treatment. This is based on the fact that each 5-point change on the PHQ-9 corresponds with a moderate effect size on multiple domains of health-related quality of life and functional status.[9] Also, an absolute PHQ-9 score of less than 10 qualifies as a partial response and a score of less than 5 as remission. These numbers are obviously simple rules of thumb that require clinical evaluation of the individual patient.

PHQ-8: AN ALTERNATIVE DEPRESSION SEVERITY MEASURE FOR SOME TYPES OF RESEARCH

Because the PHQ-9 has been increasingly used in clinical research, there have been certain types of projects in which omitting the ninth item inquiring about "thoughts that you would be better off dead or of hurting yourself in some way" is desirable. These include population or clinical samples in which one or more of the following three criteria are met: (1) the risk of suicidality is felt to be extremely low or negligible; (2) depression is being assessed as a secondary outcome in studies of other medical conditions; and/or (3) data is being gathered in a self-administered fashion rather than by direct interview, such that further probing about positive responses to item nine is not feasible. Examples include mailed questionnaires, telephone-administered interactive voice recording, or Internet surveys.

Therefore, we analyzed data from the original PHQ studies to determine the operating characteristics of the PHQ-8 (ie, all items on the PHQ-9 scale except the ninth item). The PHQ depression module classifies patients into three groups: major depression, other depression (which includes patients with both dysthymia and minor depression), and no depression. First, we compared the PHQ-8 and PHQ-9 in their ability to predict any depressive disorder (ie, either major depression or other depression). As shown in Table 2, there is a similar likelihood of any depressive disorder on the PHQ-8 and PHQ-9 at each level of depression severity level.

Second, we focused on major depression, and compared the sensitivity, specificity, and positive predictive value of the PHQ-8 and PHQ-9 across a range of cutpoints that were examined in the original PHQ-9 paper.[9] Again, as shown in Table 3, the PHQ-8 and PHQ-9 had similar operating characteristics, regardless of the cutpoint.

The reason that deletion of the ninth item has only a minor effect on the actual PHQ-9 score is that thoughts of death or self-harm are typically less common in a primary care depressed population than in the more severely depressed patients referred to a mental health specialist. Also, even patients who endorse this item often do so at a low threshold (eg, "several days"). Thus, this item typically contributes, on average, only a point or two to the overall PHQ score. In using the PHQ clinically, it is obviously essential to include this ninth item so that those patients endorsing it can be further questioned about suicidal ideation. However, even in primary care

patients depressed enough to warrant antidepressant therapy, few of those endorsing this ninth item actually have true suicidal ideation when further probed about the meaning of their response.[21] Still, because nearly half of suicide victims have contact with a primary care provider within 1 month of suicide, the PHQ-9 should be the measure of choice in most instances where the aim is to evaluate clinical populations for depression.[22] However, the PHQ-8 may be an acceptable alternative to the PHQ-9 in certain research studies that meet one of the three criteria initially outlined above.

PHQ-2: A TWO-ITEM MEASURE FOR DEPRESSION SCREENING

In some cases, the purpose is simply to screen for depression, not to assess depression severity. Even briefer versions may be desirable where the aim is to include just a few depression questions in multi-purpose health questionnaires. The US Preventive Services Task Force recently recommended depression screening as part of routine care.[23] However, brevity is essential to accomplish this in the busy general medical setting where patient volume is high, most visits are brief, and depression is simply one of many conditions that the primary care clinician is responsible for recognizing and managing.[24-26] Previous studies have suggested that one or two questions about depressed mood and, possibly, anhedonia are quite sensitive as a first-stage depression screening procedure.[1,2,27]

Therefore, we examined the performance of the PHQ-2, (ie, the first two items of the PHQ-9 that inquire about depressed mood and anhedonia). The PHQ-2 score can range from 0 to 6. Analyzing data from the PHQ Primary Care study of 3000 patients,[7] the optimal cutpoint as a depression screener turned out to be 3. A PHQ-2 score of 3 or greater has a sensitivity for major depression of 83%, a specificity of 90%, and a positive likelihood ratio of 2.9. Lowering the cutpoint to 2 increases sensitivity to 93% but reduces substantially the specificity to 74% and the positive likelihood ratio to 0.6. This would produce an unacceptably high rate of false positives: almost 80% of primary care patients with a PHQ-2 score of 2 or greater do not have major depression.

CONCLUSION

A number of comparable measures exist for detecting depression[1,2,4,11,27] including multiple self-administered scales. In contrast, it is less clear what the optimal measure for monitoring response to treatment may be, especially outside the setting of a clinical trial. Sensitivity to change is clearly a necessary feature, but other pragmatic considerations include the number of items, time required for completion, mode of administration (self-rating versus interviewer-administered scale), complexity of scoring, inter-rater agreement, and special training requirements. The specific items included in the scale are another factor. One advantage of the PHQ-9 is its exclusive focus on the 9 diagnostic criteria for *DSM-IV* depressive disorders. On the other hand, some may argue that instruments including symptoms not in the *DSM-IV* criteria (eg, loneliness, hopelessness, anxiety) may have additional clinical value. At the same time, it is possible that such scales are less specific for major depression and

TABLE 3

Comparison of the Operating Characteristics of PHQ-8 versus PHQ-9 in Diagnosing Major Depression in 3000 Primary Care Patients

PHQ Cutpoint	Sensitivity		Specificity		Positive Predictive Value	
	PHQ-8	PHQ-9	PHQ-8	PHQ-9	PHQ-8	PHQ-9
⩾ 10	99+	99+	92	91	57	55
⩾ 11	99	99	95	94	66	64
⩾ 12	97	98	97	96	75	73
⩾ 13	94	97	98	97	83	80
⩾ 14	87	92	99	99	90	88
⩾ 15	81	86	99	99+	94	93

Nine Symptom Checklist

Over the last 2 weeks, how often have you been bothered by any of the following problems?

	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things.....	0	1	2	3
2. Feeling down, depressed, or hopeless.....	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much.....	0	1	2	3
4. Feeling tired or having little energy.....	0	1	2	3
5. Poor appetite or overeating.....	0	1	2	3
6. Feeling bad about yourself - or that you are a failure or have let yourself or your family down.....	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television.....	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual.....	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way.....	0	1	2	3

(For office coding: Total Score _____ = ___ + ___ + ___)

If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all	Somewhat difficult	Very difficult	Extremely difficult
☐	☐	☐	☐

From the Primary Care Evaluation of Mental Disorders Patient Health Questionnaire (PRIME-MD PHQ). The PHQ was developed by Drs. Robert L. Spitzer, Janet B.W. Williams, Kurt Kroenke and colleagues. For research information, contact Dr. Spitzer at rls8@columbia.edu. PRIME-MD® is a trademark of Pfizer Inc. Copyright© 1999 Pfizer Inc. All rights reserved. Reproduced with permission.

other mood disorders and may discriminate depression from anxiety or even general psychological distress with less accuracy.

Detecting depression and initiating treatment are necessary but often insufficient steps to improve outcomes in primary care.[28] Monitoring clinical response to therapy is also critical. Multiple studies have shown that monitoring is often inadequate, resulting in clinician failure to detect medication noncompliance, increase the antidepressant dosage, change or augment pharmacotherapy, or add psychotherapy as needed.[28,29] Having a simple self-administered measure to complete either in the clinic or by telephone administration (eg, nurse administration[30] or interactive voice recording[31]) would provide an efficient means to assess the number and severity of the nine *DSM-IV* symptoms. Also, physicians prefer to quantify a disorder when possible, like a blood pressure reading in hypertensive patients or an electrocardiographic tracing in patients with heart disease. The PHQ-9 might be considered a type of lab test. Like blood glucose readings that serve as an entry point for diabetic patients and clinicians to communicate about disease control and to adjust therapy, PHQ-9 scores might serve a similar purpose for depressed patients and their physicians.

Brevity coupled with its construct and criterion validity makes the PHQ-9 an attractive, dual-purpose instrument for making diagnoses and assessing severity of depressive disorders, particularly in the busy setting of clinical practice. If our preliminary data on sensitivity to change of the PHQ-9 is substantiated in several large ongoing clinical trials, it could also prove to be a useful measure for monitoring outcomes of depression therapy. Finally, alternative versions may occasionally be considered for use in certain types of research (PHQ-8) or when just a few screening items are desired (PHQ-2).

REFERENCES

1. Mulrow CD, Williams JW, Gerety MB, Ramirez G, Montiel OM, Kerber C. Case-finding instruments for depression in primary care settings. *Ann Intern Med.* 1995;122:913-921.
2. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression: two questions are as good as many. *J Gen Intern Med.* 1997;12:439-445.
3. Keller MB, Kocsis JH, Thase ME, Gelenberg AJ, Rush AJ, Koran L, et al. Maintenance phase efficacy of sertraline for chronic depression: a randomized controlled trial. *JAMA.* 1998; 280:1665-1672.
4. McDowell I, Kristjansson E, Newell C. Depression. In: McDowell I, Newell C, eds. *Measuring Health: A Guide to Rating Scales and Questionnaires*. New York: Oxford University Press, 1996:238-286.
5. Spitzer RL, Williams JBW, Kroenke K, Linzer M, deGruy FV, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. *JAMA.* 1994; 272:1749-1756.
6. Hahn SR, Kroenke K, Williams JBW, Spitzer RL. Evaluation of mental disorders with the PRIME-MD. In: Maruish M, ed. *Handbook of Psychological Assessment in Primary Care Settings*. Mahwah, NJ: Lawrence Erlbaum; 2000:191-253.
7. Spitzer RL, Kroenke K, Williams JBW, and the Patient Health Questionnaire Study Group. Validity and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study. *JAMA*.1999;282:1737-1744.
8. Spitzer RL, Williams JBW, Kroenke K. Validity and utility of the Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: the PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *Am J Obstet Gynecol*. 2000;183:759-769.
9. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16:606-613.
10. Berwick DM, Murphy JM, Goldman PA, Ware JE, Barsky AJ, Weinstein MC. Performance of a five-item mental health screening test. *Med Care*. 1991;29:169-176.
11. Pasacreta JV. Measuring depression. In: Frank-Stromborg M, Olsen SJ, eds. *Instruments for Clinical Health-Care Research*. Sudbury, MA: Jones and Bartlett Publishers, 1997: 342-360.
12. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979; 134:382-389.
13. Davidson J, Turnbull CD, Strickland R, et al. The Montgomery-Asberg Depression Scale: reliability and validity. *Acta Psychiatr Scand*. 1986;73:544-548.
14. Lambert MJ, Hatch DR, Kingston MD, et al. Zung, Beck, and Hamilton rating scales as measures of treatment outcome: a meta-analytic comparison. *J Consult Clin Psychology.* 1986;54:54-59.
15. Katon W, Robinson P, Von Korff M, et al. A multifaceted intervention to improve treatment of depression in primary care. *Arch Gen Psychiatry.* 1996;53:924-932.
16. Katon W, Von Korff M, Lin E, et al. Collaborative management to achieve treatment guidelines: impact on depression in primary care. *JAMA*. 1995;273:1026-1031.
17. Williams JW Jr, Barrett J, Oxman T, Frank E, Katon W, Sullivan M, et al. Treatment of dysthymia and minor depression in primary care: a randomized controlled trial in older adults. *JAMA*. 2000;284:1519-1526.
18. Unutzer J, Williams JW Jr, Callahan CM, et al. Improving primary care for depression in late life: the design of a multicenter randomized trial. *Med Care*. 2001;39:785-799.
19. Simon GE, Revicki D, Von Korff M. Telephone assessment of depression severity. *J Psychiatr Res*. 1993;27:247-252.
20. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care.* 1989;27:S178-S189.
21. Pendergast KB, West SL, Wilson AE, Swindle R, Kroenke K. Development and use of a suicidal assessment algorithm for telephone interviewers. Pharmacoepidemiol *Drug Safety*. 2000;9(suppl 1):101.
22. Luoma JB, Martin CE, Pearson JL. Contact with mental health and primary care providers before suicide: a review of the evidence. *Am J Psychiatry*. 2002;159:909-916.
23. US Preventive Services Task Force. Screening for depression: recommendations and rationale. *Ann Intern Med.* 2002;136:760-764.
24. Klinkman MS. Competing demands in psychosocial care: a model for the identification and treatment of depressive disorders in primary care. *Gen Hosp Psychiatry*. 1997;19:98-111.
25. Williams JW Jr. Competing demands: does care for depression fit in primary care? *J Gen Intern Med*. 1998;13:137-139.
26. Kroenke K. Discovering depression in medical patients: reasonable expectations. *Ann Intern Med*. 1997;126:463-465.
27. Williams JW Jr., Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed? *JAMA*. 2002;287:1160-1170.
28. Kroenke K, Taylor-Vaisey A, Dietrich AJ, Oxman TE. Interventions to improve provider diagnosis and treatment of mental disorders in primary care: a critical review of the literature. *Psychosomatics*. 2000;41:39-52.
29. Simon GE. Can depression be managed appropriately in primary care? *J Clin Psychiatry*. 1998;59 (suppl 2):3-8.
30. Hunkeler EM, Meresman J, Hargreaves WA, et al. Efficacy of nurse telehealth care and peer support in augmenting treatment of depression in primary care. *Arch Fam Med*. 2000;9:700-708.
31. Kobak KA, Taylor LvH, Dottl SL, Greist JH, Jefferson JW, Burroughs D, et al. A computer-administered telephone interview to identify mental disorders. *JAMA*.